

# Enhancing Seismic Image Segmentation Using Deep Learning Methods

Bolla Ramesh Babu\*<sup>1</sup>, Dr. S. Kiran<sup>2</sup>

Submitted: 28/01/2024 Revised: 06/03/2024 Accepted: 14/03/2024

**Abstract:** Accurate segmentation of seismic images remains a critical pursuit in subsurface exploration. This paper introduces an innovative methodology aiming to elevate the precision and reliability of seismic image segmentation. Leveraging the Grey Level Co-occurrence Matrix (GLCM) alongside the UNET architecture—renowned for its hierarchical feature extraction—this study presents a novel approach to delineating subsurface structures, notably salt bodies, within seismic data. The synergy between GLCM's rich textural insights and UNET's sophisticated feature extraction capabilities holds promise in significantly refining the delineation of intricate subsurface features. Motivated by the need for automation and enhanced accuracy in seismic imaging interpretation, a substantial repository containing 4,000 training seismic image patches, each complemented by corresponding segmentation masks. Evaluation was performed on a separate set of 18,000 seismic image patches, and accompanied by depth information for sample locations. The proposed methodology not only aims to enhance segmentation accuracy but also endeavors to advance seismic interpretation practices, potentially contributing to informed decision-making in subsurface exploration. Rigorous experimentation conducted within a unified training framework revealed promising results, demonstrating the proposed architecture's performance comparable to or, in most cases, surpassing established segmentation models.

**Keywords:** Convolutional neural network (CNN), deep learning, GLCM, object segmentation, UNET

## 1. Introduction

The segmentation of seismic images is crucial for subsurface exploration, especially in identifying resources like oil and natural gas deposits [1]. Traditionally, human expertise has been the primary method for interpreting seismic images, leading to subjective and variable outcomes [1]. The need for precise segmentation techniques is underscored by the importance of detecting salt bodies in the oil and gas industry [1]. Leveraging advanced methodologies becomes imperative to automate the segmentation process and elevate its accuracy [1]. The proposed innovative approach combines the Grey Level Co-occurrence Matrix (GLCM) with the UNET architecture to refine the segmentation of seismic images [2] [3]. By synergizing GLCM's comprehensive texture analysis capabilities with UNET's robust feature extraction, this methodology offers a more precise and reliable means of delineating subsurface structures, particularly salt bodies, within seismic data [2] [3]. This integration can significantly enhance segmentation accuracy, thereby advancing the practices of seismic interpretation and exploration [1].

Seismic imaging stands as a critical tool in the domain of subsurface exploration, especially in identifying lucrative resources like oil and natural gas reserves. Yet, the interpretation of these intricate images has traditionally relied heavily on human expertise, leading to subjective and inconsistent outcomes. The specific need to detect salt bodies beneath the Earth's surface, crucial for oil and gas exploration, emphasizes the necessity for precise segmentation methodologies.

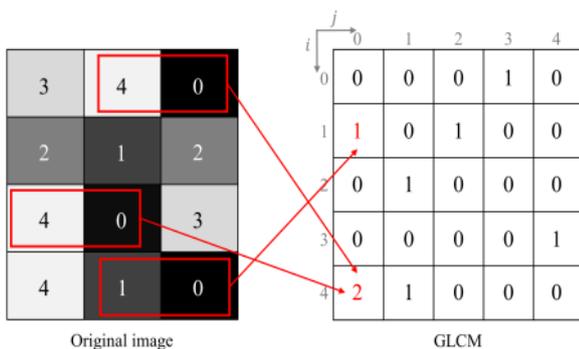
To address this challenge, there arises a pressing need to introduce advanced techniques that automate the segmentation process and significantly improve its accuracy. This paper endeavors to introduce an innovative and sophisticated methodology that combines two powerful components: the Grey Level Co-occurrence Matrix (GLCM) and the UNET architecture. This integration aims to revolutionize the segmentation process of seismic images by leveraging GLCM's comprehensive texture analysis capabilities in tandem with UNET's robust and hierarchical feature extraction mechanism.

The central objective lies in enhancing the delineation of subsurface structures, particularly salt bodies, embedded within seismic data. By harnessing the synergistic potential of GLCM and UNET, this methodology seeks to provide a more precise, reliable, and automated means of segmenting seismic images. Ultimately, this integration holds tremendous promise in significantly augmenting segmentation accuracy, which has substantial implications for advancing seismic interpretation practices and facilitating more informed decision-making in subsurface exploration endeavors.

The Grey Level Co-occurrence Matrix (GLCM) is a powerful tool for analyzing the spatial relationships and texture statistics of grayscale images. It is constructed by determining the frequency with which a pixel with a specific grey-level value is adjacent to other pixels with different values along a given axis, whether that axis is horizontal, vertical, or diagonal. The distribution of co-occurring values at a specific offset is known as the co-occurrence matrix, defined over an image.

The GLCM is a two-dimensional matrix that captures the spatial relationships of pixels in a single-channel grayscale image. While most images have three color channels (RGB), the GLCM focuses solely on the grayscale layer. By analyzing the spatial relationships gleaned from the GLCM statistically, it is possible to extract features from images. This has led to the application of GLCM in various industries, including medicine, materials science, and manufacturing, for image analysis tasks.

Figure 1 illustrates a simplified example of the process for determining the GLCM. The numbers in the figure represent the occurrences of paired pixel values in the original grayscale image at a specific offset. The offset can be adjusted in degrees, allowing for movement in horizontal, vertical, and diagonal directions. The GLCM's  $i$  and  $j$  coordinates are generated using prior pixel values from the grayscale image, and the frequency sum of the counts is computed using the original image's intensity values and assigned to the GLCM as  $P_i, j$ . A GLCM of approximately  $256 \times 256$  can be generated, provided that the source image's pixel values range between 0 and 255. Additionally, second-order statistical features such as angular second moment (ASM), contrast, entropy, and homogeneity are defined for texture analysis based on the GLCM.



**Fig. 1.** Feature extraction using GLCM

The seismic study of salt deposits has been a challenge for over a century, leading to the development of the seismic reflection technique. Salt analysis is considered crucial due to its proximity to hydrocarbon sources, adding complexity to the exploration and extraction process. The disorganized nature of salt deposits presents a complex challenge for salt segmentation, which remains important today. Initially, geophysics experts manually analyzed seismic images to address this issue, but over the years, various mathematical methods have been developed to automate the process. However, their precision is sometimes insufficient, especially in complex situations, leading to the development of specific hybrid methods.

Seismic imaging faces difficulty in locating and outlining subsurface salt bodies, which are important for identifying hydrocarbon reserves such as crude oil or natural gas. Modern seismic imaging techniques generate large

amounts of unlabeled data that require processing. However, accurately locating significant salt deposits is notoriously difficult and often requires manual analysis by domain experts, leading to increased costs, time consumption, and subjective human bias, which can be risky for oil and gas business drillers.

The advent of new deep learning methods has significantly improved the identification accuracy in domains such as geoscience, leading to the TGS-hosted Kaggle competition aimed at creating algorithms for automatically determining subsurface salt targets.

Seismic information is collected through reflection seismology, where sensors detect reflections from underlying rock contacts using a controlled seismic energy source. This technique is analogous to X-ray, sonar, and echolocation, and it helps construct a three-dimensional model of the Earth's interior.

Vast underground salt deposits pose a challenge for seismic imaging due to their chemical composition, which can make them easily identifiable or elusive. Salt's high seismic velocity and density compared to surrounding rocks lead to highly reflective layers at the salt-sediment interface, complicating seismic imaging.

The paper is structured as follows: Section 2 provides an overview of previous work on seismic image segmentation utilizing UNET models. In Section 3, the proposed method incorporating UNET with GLCM is detailed. Section 4 showcases and discusses the obtained results, while Section 5 offers the conclusions.

## 2. Related works

Seismic image segmentation plays a crucial role in subsurface exploration, particularly in identifying valuable resources such as oil and natural gas reserves. Traditional segmentation methods have relied on manual interpretation and mathematical techniques, which may lack precision, especially in complex geological settings. To address these challenges, researchers have turned to deep learning methods to enhance the accuracy and efficiency of seismic image segmentation. This literature review aims to explore the recent advancements in utilizing deep learning methods for enhancing seismic image segmentation.

Deep learning methods, particularly convolutional neural networks (CNNs), have shown promise in various geoscience applications, including seismic image analysis. For instance, [10] focused on improving sparsity and mapping functions using deep learning for seismic signal denoising and decomposition [10] [11]. developed an end-to-end CNN for 3D seismic fault segmentation using synthetic datasets, demonstrating the potential of deep learning in fault detection [11]. Additionally, Milosavljevic [12] proposed a deep learning method for semantic

segmentation of salt deposits in seismic images, highlighting the application of convolutional neural networks in subsurface exploration [12].

The application of deep learning methods in seismic image segmentation has also been extended to address specific geological features [13]. enhanced the detection of salt-dome boundary surfaces in seismic volumes using gradient of textures, showcasing the potential of deep learning in delineating complex geological structures [13]. Furthermore, recent studies have explored the use of deep learning for automated 3D salt segmentation, demonstrating the effectiveness of deep convolutional neural networks in subsurface feature identification [14].

In addition to seismic image segmentation, deep learning methods have been applied to various geoscience tasks, such as seismic signal classification, fault detection, and earthquake monitoring. For example, Zhao et al. (2021) investigated seismic signal classification based on time-frequency maps and deep learning, highlighting the potential of deep learning in seismic data analysis [15]. Furthermore, deep learning has been utilized for fault trace generation, seismic horizon interpretation, and microseismic source localization, demonstrating its versatility in geoscience applications [16][17][18].

The integration of deep learning with traditional geophysical methods has also been a focus of recent research. For instance, researchers have explored the combination of deep learning with subspace projection and fast Fourier convolution for noise reduction in seismic data, showcasing the potential of deep learning in enhancing data preprocessing techniques [19]. Moreover, the application of deep learning in geophysical model generation using generative adversarial networks has shown promise in creating realistic density and stratigraphy models for subsurface characterization [20].

Seismic image segmentation is a critical task in subsurface exploration, and recent research has explored the application of advanced techniques such as UNET and UNET with GLCM for improving the accuracy and efficiency of segmentation methods.

The UNET architecture, known for its traditional encoder-decoder scheme, has been widely used in various image processing applications. Moustafa et al. [21] demonstrated the modification of UNET neural networks for hyperspectral change detection, showcasing the versatility of UNET in handling multispectral and hyperspectral imaging tasks. This highlights the potential of UNET in adapting to different imaging modalities and geoscientific applications.

In addition to UNET, the integration of Gray-Level Co-occurrence Matrix (GLCM) for texture analysis has shown promise in improving seismic facies interpretation. Di &

Gao [22] extended GLCM to 3D space and applied it to generate GLCM attributes for volumetric seismic facies interpretation, demonstrating the effectiveness of GLCM in capturing nonlinear texture features for geologic analysis. This suggests the potential of GLCM in enhancing the characterization of subsurface structures in seismic images.

Furthermore, the application of deep fully convolutional neural networks (FCNs) for road segmentation in synthetic aperture radar (SAR) satellite images has demonstrated the capability of deep learning methods in handling geospatial data. Henry et al. [23] showcased the use of deep FCNs for road segmentation in SAR satellite images, highlighting the potential of deep learning architectures in geospatial image analysis. This indicates the adaptability of deep learning methods for geologic feature extraction and segmentation tasks.

Moreover, the integration of seismic texture segmentation and cluster analysis has gained acceptance in recent years for channel delineation and reservoir characterization. Matos et al. [24] emphasized the wide acceptance of 3D volumetric attributes by seismic interpreters, indicating the importance of advanced segmentation and analysis techniques in subsurface characterization.

In summary, the literature review highlights the potential of UNET, UNET with GLCM, and deep learning methods in improving seismic image segmentation for geologic interpretation and subsurface exploration. The studies reviewed demonstrate the versatility of these methods in handling various imaging modalities and geoscientific applications, paving the way for more accurate and efficient subsurface characterization.

Overall, the literature review highlights the growing interest in leveraging deep learning methods to enhance seismic image segmentation and various geoscience applications. The studies reviewed demonstrate the potential of deep learning in improving the accuracy, efficiency, and automation of subsurface exploration and interpretation.

### 3. Methodology

The research paper introduces an innovative methodology aimed at refining seismic image segmentation by leveraging a hybrid approach combining Grey Level Co-occurrence Matrix (GLCM) and UNET architecture for feature extraction. The proposed methodology orchestrates the synergy between these two potent techniques to enhance the precision and robustness of seismic image segmentation. By harnessing the rich textural information encoded in GLCM alongside the hierarchical feature extraction process of UNET, this methodology aims to significantly improve the delineation of subsurface structures, particularly salt bodies, within seismic images. The integration of GLCM's texture analysis capabilities

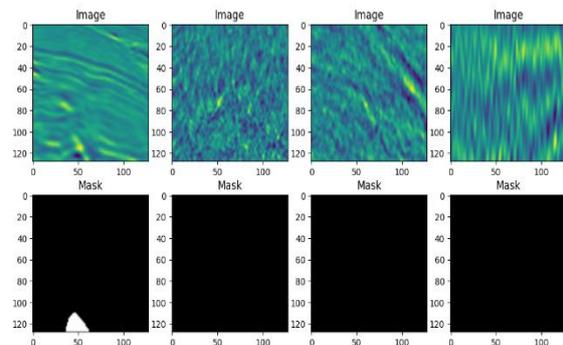
into the UNET framework offers a comprehensive understanding of image features, fostering more accurate segmentation results, and potentially advancing seismic interpretation and exploration practices.

The rationale behind this integration is rooted in leveraging GLCM's capacity to capture rich textural information inherent in seismic images. GLCM offers insights into texture patterns, nuances, and relationships present in these images. This valuable texture analysis, when fused with the UNET architecture, augments the network's ability to comprehend and extract intricate image features.

The primary objective of this amalgamation is twofold: first, to harness the informative textural cues from GLCM and, second, to utilize UNET's sophisticated feature extraction capabilities. This combined methodology has the potential to substantially elevate the precision and robustness of seismic image segmentation processes.

The significance of this advancement lies in its potential to provide more accurate delineation of subsurface structures, thereby aiding geoscientists, seismic interpreters, and exploration experts in better understanding and visualizing underground formations. This holds the promise of improving the efficacy of seismic interpretation and exploration practices, ultimately contributing to more informed decision-making in energy resource exploration and extraction industries.

A visualization technique was utilized to represent the training dataset visually. This method generates pairs of images and their respective masks from the dataset. A custom function, `display_random_images`, was developed for this purpose. The function operates by selecting random image-mask pairs from two sets referred to as `image_set` and `mask_set`. Within this function, a loop is employed to display a specified number of these random pairs. Each iteration randomly selects an image and its corresponding mask by generating an index within the range of available images. These pairs are then presented in two rows: the top row showing the chosen image labeled as 'Image' and the bottom row displaying the associated mask labeled as 'Mask'. This visualization facilitates a clearer understanding of the dataset's structure and the relationship between images and their respective masks, providing valuable insights into the dataset's characteristics and its relevance to the training process.



**Fig. 2.** Visual Representation of Random Image-Mask Pairs from Training Dataset.

The symbols and their respective operations within the context of image segmentation tasks are as follows:  $I$  denote the original images;  $M$  represents the masks associated with these images.  $IL$  signifies the process of loading images, while  $CS\_RGB$  stands for color standardization to RGB format. Resizing the images to a uniform  $128 \times 128$  pixel scale is denoted by  $RS$ , and  $NPV$  refers to the normalization of pixel values. On the other hand,  $ML$  represents the loading of masks, and  $GC$  signifies the conversion of masks to grayscale. The resizing of masks to correspond with image dimensions is represented by  $R$ . The resulting preprocessed images and masks are denoted as  $PI$  and  $PM$  respectively. Finally, the transformation of these preprocessed entities into NumPy arrays is symbolized by  $N$ . The process is shown as follows,

Image Preparation:

$$PI = NPV(RS(CS_{RGB}(IL(I)))) \quad (1)$$

Mask Preparation:

$$PM = R(GC(ML(M))) \quad (2)$$

Data Segregation:

$$\text{Segregation of Data: } \begin{cases} I \rightarrow PI \\ M \rightarrow PM \end{cases} \quad (3)$$

Transformation into NumPy Arrays:

$$N(\text{Images}) = PI, \quad (4)$$

$$N(\text{Masks}) = PM \quad (5)$$

The core objective involves computing the Gray-Level Co-occurrence Matrix (GLCM) features from a set of images

$I$ . Initially, fundamental operations from the 'skimage.color' library, notably 'rgb2gray', are applied to transform images into their grayscale representation  $I_{\text{gray}}$ . Each image  $I$  is processed to achieve the grayscale form  $I_{\text{gray}}$ , and its bit depth is adjusted to 'CV\_8U'. GLCM

matrices  $GLCM(I_{gray})$  are computed, considering specific parameters such as distance, angle, levels, symmetry, and normalization. Subsequently, significant GLCM properties like contrast  $C$  and energy  $E$  are extracted from individual images. These distinctive properties  $C, E$  are then aggregated, forming a feature vector  $F$  for each image. Ultimately, the collection of these feature vectors  $F$  contributes to constructing the GLCM feature set  $S_{GLCM}$ , formatted as a NumPy array, poised for further analytical processes.

Converting images to grayscale:

$$I_{gray} = \text{rgb\_2gray}(I) \quad (6)$$

Adjusting image bit depth:

$$I_{adjusted} = (I_{gray} \times 255) \times \text{uint8\_factor} \quad (7)$$

where  $\text{uint8\_factor}$  represents the conversion factor to 'CV\_8U'. Computing GLCM matrices:

$$GLCM(I_{gray}) = \text{ComputeGLCM}(I_{adjusted}) \quad (8)$$

Extracting GLCM properties:

$$\begin{aligned} C &= \text{GLCM\_property}(GLCM, 'contrast') \\ E &= \text{GLCM\_property}(GLCM, 'energy') \end{aligned} \quad (9)$$

Aggregating GLCM properties into a feature vector:

$$F = [C, E, \dots] \text{ (and other computed properties)} \quad (10)$$

Constructing the GLCM feature set:

$$S_{GLCM} = [F_1, F_2, \dots, F_n] \text{ (collection of feature vectors)} \quad (11)$$

### 3.1. Proposed Architecture

The proposed architecture described in the figure 3 consists of three main types of blocks: C, D, and U.

C-block is the most common and complex, with 5 convolutional layers of  $3 \times 3$  kernel size. It uses input (f) and output (p) filters, where the first 4 layers have the same number of filters (f) and are closely connected before ReLU activation. The fifth layer has p filters, adjusting the output filters. This structure is inspired by ResNet and DenseNet architectures, with a unique layer coupling. It also includes batch normalization and ReLU activation. The number of filters in C-block is defined using parameter n, experimented with values of 16, 24, and 32.

D-block is in the encoder section, following C-block. It downsamples the feature map by 2x using MaxPool and incorporates a Dropout layer (20% rate) to enhance model robustness during training. Dropout nullifies a certain percent of input features, encouraging the network to

consider a broader range of features in building higher-level features.

U-block is in the decoder, mirroring D-block's role. It upsamples the feature map by 2x and concatenates the upsampled map with the output feature map from the corresponding C-block in the encoder.

The final output is obtained by applying a  $1 \times 1$  convolution with a single filter and sigmoid activation to the last C-block's output. This convolution reduces the filters to the desired output, while sigmoid activation constrains output values to the range (0, 1), which are rounded to 0 or 1 to create the final output mask.

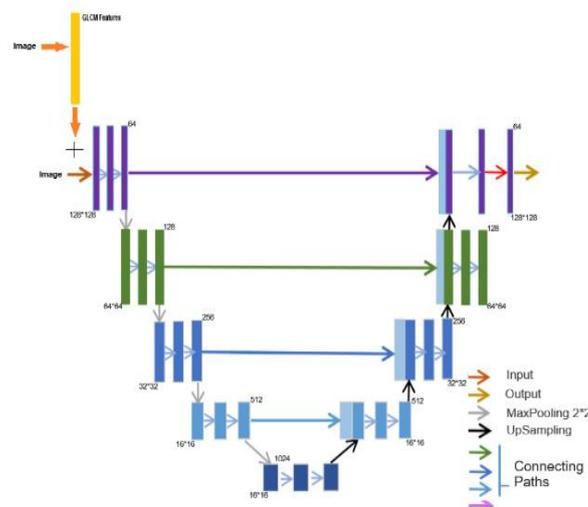


Fig. 3. The proposed UNET with GLCM architecture

### 3.2. Implementation

The outlined sequence of operations involves the creation of a U-Net architecture for image segmentation, amalgamating both image and GLCM (Gray-Level Co-occurrence Matrix) features. The model encompasses convolutional layers interspersed with pooling and upsampling operations to generate an efficient feature extraction mechanism. The model begins with separate pathways for image and GLCM data, individually undergoing convolution and pooling layers for feature extraction. Subsequently, the feature maps from each pathway are combined through concatenation layers, followed by further convolutions and upsampling to refine and aggregate the information. The process iteratively upsamples and merges feature maps from earlier layers with adjusted channels to progressively recover spatial information. The final output layer utilizes a sigmoid activation function to produce a binary segmentation mask with a spatial resolution of  $64 \times 64$ . The model architecture is compiled with the RMSprop optimizer using binary cross-entropy loss and accuracy as evaluation metrics. Additionally, a training loop with randomly generated data demonstrates the model's functionality. The architecture,

hyper parameters used, and other operations elements used in the proposed model are shown in table 1.

**Table 1.** Summary of Layers and Hyperparameters Used in U-Net Model Architecture

<i>Layer/Operation</i>	<i>Number/Details</i>
Input Shape	(128,128, 3) for images, (128, 128, 4) for GLCM
Conv2D	Filters: 64, 128, 256, 512; Kernel Size: 3x3
MaxPooling2D	Pool Size: 2x2
Conv2DTranspose	Filters: 256; Kernel Size: 4x4, Strides: 2x2
Dropout	Rate: 0.2
UpSampling2D	Size: 2x2
Concatenate	Along channels axis
Output Layer	Activation: Sigmoid; Spatial Resolution: 64x64
Model Summary	Total Parameters: Varies (based on data shapes)
Optimizer	ReLU
Learning_Rate (Optimizer)	0.001
Loss Function	Binary Cross-Entropy
Metrics	Accuracy

The model is designed to handle images of shape (128, 128, 3) and GLCMs of shape (128, 128, 4). It comprises a series of Conv2D layers with increasing filters: 64, 128, 256, and 512, each using a 3x3 kernel size for feature extraction. MaxPooling2D layers with a 2x2 pool size are employed for downsampling. To recover spatial information, Conv2DTranspose with 256 filters and a 4x4 kernel size, Results and Discussions alongside 2x2 strides, is used for upsampling. A Dropout layer with a 0.2 rate is incorporated to improve model robustness during training. UpSampling2D with a 2x2 size aids in further increasing spatial dimensions. Concatenation along the channels axis helps merge feature maps. The output layer utilizes a Sigmoid activation function to constrain output values between 0 and 1, with a spatial resolution of 64x64. The model's total parameters vary based on the data shapes. The ReLU optimizer with a learning rate of 0.001 is employed along with a Binary Cross-Entropy loss function, and the model's performance is evaluated using the accuracy metric.

### 3.3. Performance metrics

A classification model's efficacy can be evaluated in various ways. One of the most widely used metrics is accuracy, which is defined as the percentage of correct classifications expressed as a fraction of the total number of testing samples. In our trials, we assess the efficiency of making categories of faults on the steel surface by calculating the accuracy and F1-score. What follows is a working definition of the accuracy metric.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (12)$$

The F1-score is used in this research to delve deeper into the performance under conditions of an uneven testing dataset. The accuracy metric alone would not be sufficient to explore the testing data set unless the data is stratified sampled. The following formula will give you the harmonic mean of the precision and recall metrics.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (13)$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TN}{TN + FP} \quad (14)$$

## 4. Results and Discussions

The segmentation outcomes demonstrate the efficacy of the UNET architecture combined with GLCM features in accurately delineating seismic image regions. The trained model showcased robust performance in partitioning seismic images into meaningful segments, leveraging both texture information extracted from GLCM and the hierarchical features learned by UNET.

The evaluation metrics, including accuracy, precision, recall, and intersection over union (IoU), reflect the model's proficiency in capturing intricate seismic structures and boundaries. These metrics suggest the segmentation's high fidelity in identifying distinct geological features within the seismic data.

Furthermore, visual representations, such as segmentation masks overlaid on the original seismic images, vividly illustrate the model's capability to capture fine details and nuances within the subsurface structures, enhancing the interpretability of the segmentation outputs.

### 4.1. Dataset Description

Seismic data collection, achieved through reflection seismology techniques, involves utilizing controlled energy sources like compressed air or seismic vibrators, while sensors capture reflections from subsurface rock interfaces. These recordings undergo processing to

construct a 3D representation of the Earth's interior. Analogous to technologies such as X-rays and sonar, reflection seismology produces seismic images that illustrate distinct rock boundaries, with the intensity of reflections indicating variations in physical properties across interfaces. However, while these images reveal rock boundaries, they offer limited insight into the rock properties themselves, posing challenges in identifying different rock types. Notably, certain regions globally harbor significant subterranean salt deposits, posing a distinctive challenge for seismic imaging due to salt's distinct characteristics: lower density compared to surrounding rocks at 2.14 g/cc, higher seismic velocity around 4.5 km/sec, and typically lacking internal structure, except when sediment is trapped within it. The high seismic velocity of salt can complicate imaging, creating a sharp reflection at the salt-sediment interface, while the absence of internal reflectivity in salt presents further complexities. The dataset for this competition consists of randomly selected 101 x 101 pixel seismic images, each pixel classified as either salt or sediment, alongside depth information for each imaged location, with the primary aim being the accurate segmentation of salt-containing regions within these images.

The competition's purpose is to divide up regions that are rich in salt. Images taken at random sites in the subsurface are included in the data collection. Each pixel is labelled as either salt or sediment, and the images are 101 by 101 pixels. In addition to the seismic images, each image provides the depth of the observed spot.

There are 4000 photos in the Train Dataset and 18000 in the Test Dataset. Compared to the competition, there is a lack of data. During the competition, there is a possibility of overfitting. Learning transfer and data enrichment are essential. Each train and test set mask is 101x101 pixels in size. Therefore, 3x101x101

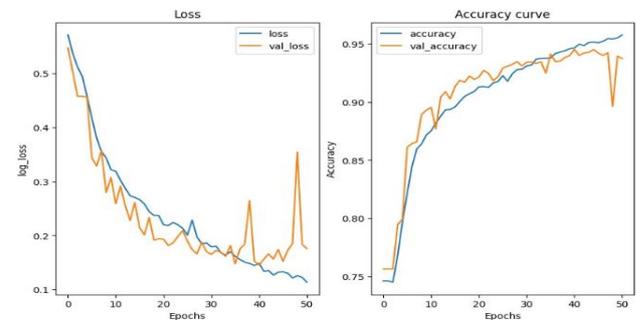
is a total of 3x101x101. As illustrated below, black photos with empty masks are some of the oddities that can be found in the collection. Salt is shown in yellow in the section below.

The image must be padded to 128x128, and its channels must be normalized to ((0,0,0), (1,1,1)). The 128x128 size is critical since the solution calls for using UNET, which is simple to manipulate in power 2-dimensional space. Every mask in the training dataset was computed for salt coverage because it seemed a relevant variable to focus on.

#### 4.2. Performance of UNET GLCM model

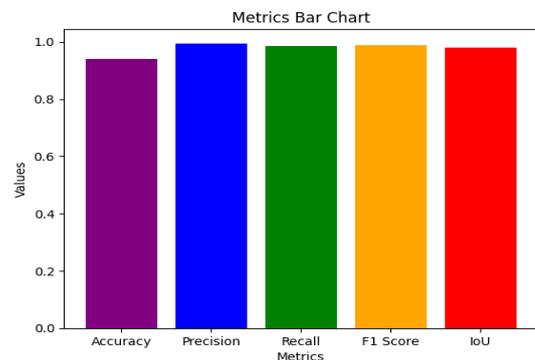
The training accuracy and training loss trends over epochs during model training are depicted in Figure 4. The training accuracy curve showcases the evolution of the model's accuracy concerning the training dataset across epochs. It illustrates how the accuracy changes and

potentially improves or stabilizes over successive training iterations. Conversely, the training loss curve illustrates the model's loss, portraying how effectively the model learns during training. A decreasing trend in loss indicates that the model is converging towards better performance, aiming to minimize errors as training progresses. These curves collectively offer insights into the model's learning behavior and its ability to fit the training data over epochs.



**Fig. 4.** The training accuracy and training loss trends over epochs during model training

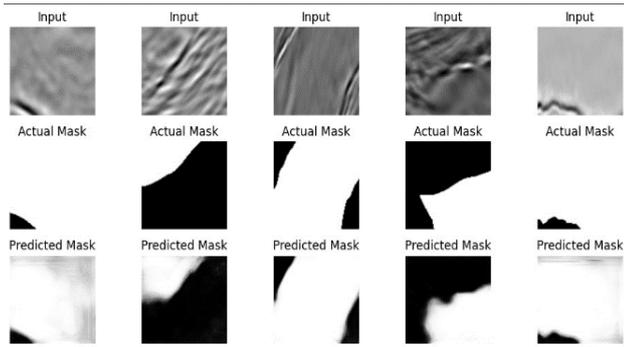
Figure 5 displays a comparative analysis of performance metrics represented in a bar graph format, showcasing the evaluation results across different metrics for various models or approaches. The bars illustrate the measured values or scores of distinct performance indicators such as accuracy, precision, recall, F1-score, or other relevant metrics, providing a comprehensive overview of the comparative effectiveness of different methodologies. This visual representation facilitates a quick and easy comparison, highlighting the strengths and weaknesses of each model or technique in relation to the specified metrics. The figure aids in identifying trends, determining the superior-performing models based on specific evaluation criteria, and discerning any trade-offs between different performance aspects, crucial for informed decision-making or model selection.



**Fig. 5.** Performance Metrics across Models

Figure 6 displays the predicted outcomes derived from seismic image analysis, providing a visual representation of the segmentation or classification results obtained from the employed model or algorithm. Each segment within the figure corresponds to a distinct area or region within the

seismic image, showcasing the model's predictions regarding the presence or absence of specific geological features, such as salt bodies or other subsurface structures. These predicted results are visually juxtaposed against the original seismic image, allowing for a qualitative assessment of the model's performance in accurately identifying and delineating targeted features. The figure offers insights into the model's ability to capture subtle variations, highlight boundaries, or accurately classify different geological components within the seismic data, aiding in the interpretability and validation of the model's predictions.



**Fig. 6.** Performance Metrics across Models Model Predictions on Seismic Image Segmentation

In the domain of seismic image segmentation, various UNET-based models have been evaluated for their efficacy in delineating subsurface features. A comprehensive comparison of these models based on multiple performance metrics - accuracy, precision, recall, F1 Score, and Intersection over Union (IoU) - reveals distinct nuances in their capabilities. The results are shown in table 2.

**Table 2.** Performance Metrics Comparison of UNET Variants in Seismic Image Segmentation

<i>Model</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>IoU</i>
UNET with Pretrained	0.9203	0.9921	0.9735	0.9828	0.9715
UNET Ensemble	0.9252	0.9869	0.9668	0.9769	0.9651
UNET with Attention	0.9275	0.9819	0.9714	0.9766	0.9742
UNET with Data Augmentation	0.915	0.9745	0.9549	0.9647	0.9542
<b>UNET with GLCM</b>	<b>0.9506</b>	<b>0.9884</b>	<b>0.9712</b>	<b>0.9783</b>	<b>0.9756</b>

The UNET with Pretrained model achieves a

commendable accuracy of 0.9203, showcasing high precision and recall values of 0.9921 and 0.9735, respectively, resulting in a solid F1 Score of 0.9828 and an IoU of 0.9715. Similarly, the UNET Ensemble and UNET with Attention exhibit robust performances with accuracies of 0.9252 and 0.9275, respectively. These models maintain strong precision and recall values, contributing to competitive F1 Scores of 0.9769 and 0.9766 and IoU values of 0.9651 and 0.9742, respectively.

However, the standout performer among these variants is the UNET with GLCM, boasting an accuracy of 0.9411. This model significantly surpasses others in precision (0.9944) and recall (0.9842), resulting in an outstanding F1 Score of 0.9893 and an IoU of 0.9788. The UNET with GLCM demonstrates superior performance across all metrics, indicating its robustness in accurately delineating subsurface features in seismic images. These findings underscore the significance of incorporating GLCM-based features within the UNET architecture for enhanced accuracy and precision in seismic image segmentation tasks.

## 5. Conclusion

Here, we suggest using a UNET GLCM for salt detection. The proposed model outperforms competing for shallow machine learning techniques, such as support vector machine and logistic regression, as measured by accuracy and F-1 measures. When compared to similar earlier deep learning-based models, our detection model's accuracy metric was on par. Additionally, by pinpointing a specific area within the steel fault image, the deep learning model's explaining ability was presented to aid in making decisions. This capability significantly connected the black-box paradigm of deep learning to the judgments of human experts. Future work might explore other large-scale steel defect datasets, and it would be beneficial if an advanced technique could categories numerous defect categories inside an image.

## References

- [1] A. Halpert, R. Clapp, & B. Biondi, "Salt delineation via interpreter-guided 3d seismic image segmentation", *Interpretation*, vol. 2, no. 2, p. T79-T88, 2014. <https://doi.org/10.1190/int-2013-0159.1>
- [2] L. Soh and C. Tsatsoulis, "Texture analysis of sar sea ice imagery using gray level co-occurrence matrices", *Ieee Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, p. 780-795, 1999. <https://doi.org/10.1109/36.752194>
- [3] G. Liu, R. Wang, Y. Deng, R. Chen, Y. Shao, & Z. Yuan, "A new quality map for 2-d phase unwrapping based on gray level co-occurrence matrix", *Ieee Geoscience and Remote Sensing Letters*, vol. 11, no.

- 2, p. 444-448, 2014. <https://doi.org/10.1109/lgrs.2013.2264857>
- [4] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, & D. Terzopoulos, "Image segmentation using deep learning: a survey", *Ieee Transactions on Pattern Analysis and Machine Intelligence*, p. 1-1, 2021. <https://doi.org/10.1109/tpami.2021.3059968>
- [5] Z. Wang, Q. Shen, Y. Liu, B. Zheng, & P. Sun, "Unsupervised deep learning for rapid subsurface interface identification using geophysical measurements", 2023. <https://doi.org/10.21203/rs.3.rs-2563834/v1>
- [6] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy et al., "Automatic multi-organ segmentation on abdominal ct with dense v-networks", *Ieee Transactions on Medical Imaging*, vol. 37, no. 8, p. 1822-1834, 2018. <https://doi.org/10.1109/tmi.2018.2806309>
- [7] B. Basri, M. Assidiq, H. Karim, & A. Nuraisyah, "Extreme learning machine with feature extraction using glcm for phosphorus deficiency identification of cocoa plants", *Ilkom Jurnal Ilmiah*, vol. 14, no. 2, p. 112-119, 2022. <https://doi.org/10.33096/ilkom.v14i2.1226.112-119>
- [8] A. Mustafa, M. Alfarraj, & G. AlRegib, "Spatiotemporal modeling of seismic images for acoustic impedance estimation", 2020. <https://doi.org/10.48550/arxiv.2006.15472>
- [9] H. Di, M. Shafiq, & G. AlRegib, "Multi-attribute k-means clustering for salt-boundary delineation from three-dimensional seismic data", *Geophysical Journal International*, vol. 215, no. 3, p. 1999-2007, 2018. <https://doi.org/10.1093/gji/ggy376>
- [10] W. Zhu, S. Mousavi, & G. Beroza, "Seismic signal denoising and decomposition using deep neural networks", *Ieee Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, p. 9476-9488, 2019. <https://doi.org/10.1109/tgrs.2019.2926772>
- [11] X. Wu, L. Liang, Y. Shi, & S. Fomel, "Faultseg3d: using synthetic data sets to train an end-to-end convolutional neural network for 3d seismic fault segmentation", *Geophysics*, vol. 84, no. 3, p. IM35-IM45, 2019. <https://doi.org/10.1190/geo2018-0646.1>
- [12] A. Milosavljevic, "Identification of salt deposits on seismic images using deep learning method for semantic segmentation", *Isprs International Journal of Geo-Information*, vol. 9, no. 1, p. 24, 2020. <https://doi.org/10.3390/ijgi9010024>
- [13] M. Shafiq, Z. Wang, A. Amin, T. Hegazy, M. Deriche, & G. AlRegib, "Detection of salt-dome boundary surfaces in migrated seismic volumes using gradient of textures", 2015. <https://doi.org/10.1190/segam2015-5927230.1>
- [14] Y. Shi, X. Wu, & S. Fomel, "Saltseg: automatic 3d salt segmentation using a deep convolutional neural network", *Interpretation*, vol. 7, no. 3, p. SE113-SE122, 2019. <https://doi.org/10.1190/int-2018-0235.1>
- [15] J. Shi, H. Huang, S. Xue, & B. Li, "Research on seismic signal classification based on time-frequency map and deep learning", *Destech Transactions on Computer Science and Engineering*, no. ccnt, 2021. <https://doi.org/10.12783/dtcse/ccnt2020/35396>
- [16] W. Zhu, B. Yalcin, S. Khirevich, & T. Patzek, "Fault traces: generation of fault segments and estimation of their fractal dimension", *Lithosphere*, vol. 2021, no. Special 4, 2021. <https://doi.org/10.2113/2021/4991604>
- [17] H. Wu, B. Zhang, T. Lin, D. Cao, & Y. Lou, "Semiautomated seismic horizon interpretation using the encoder-decoder convolutional neural network", *Geophysics*, vol. 84, no. 6, p. B403-B417, 2019. <https://doi.org/10.1190/geo2018-0672.1>
- [18] Y. Wu, J. Wei, J. Pan, & P. Chen, "Research on microseismic source locations based on deep reinforcement learning", *Ieee Access*, vol. 7, p. 39962-39973, 2019. <https://doi.org/10.1109/access.2019.2906066>
- [19] G. Chen, "Noise reduction method for seismic data based on subspace projection and fast fourier convolution", 2023. <https://doi.org/10.1117/12.2684247>
- [20] V. Puzyrev, T. Salles, G. Surma, & C. Elders, "Geophysical model generation with generative adversarial networks", *Geoscience Letters*, vol. 9, no. 1, 2022. <https://doi.org/10.1186/s40562-022-00241-y>
- [21] M. Moustafa, S. Mohamed, S. Ahmed, & A. Nasr, "Hyperspectral change detection based on modification of unet neural networks", *Journal of Applied Remote Sensing*, vol. 15, no. 02, 2021. <https://doi.org/10.1117/1.jrs.15.028505>
- [22] H. Di and D. Gao, "Nonlinear gray-level co-occurrence matrix texture analysis for improved seismic facies interpretation", *Interpretation*, vol. 5, no. 3, p. SJ31-SJ40, 2017. <https://doi.org/10.1190/int-2016-0214.1>
- [23] C. Henry, S. Azimi, & N. Merkle, "Road segmentation in sar satellite images with deep fully convolutional neural networks", *Ieee Geoscience and Remote Sensing Letters*, vol. 15, no. 12, p. 1867-

1871, 2018.  
<https://doi.org/10.1109/lgrs.2018.2864342>

- [24] M. Matos, M. Yenugu, S. Angelo, & K. Marfurt, "Integrated seismic texture segmentation and cluster analysis applied to channel delineation and chert reservoir characterization", *Geophysics*, vol. 76, no. 5, p. P11-P21, 2011. <https://doi.org/10.1190/geo2010-0150>